Probe into Markov Chain by Counting Spanning Trees/Forests

Xinyu Mao
Instructor: Yaokun Wu

January 2, 2021

Summary

In this note, I shall present two results that reveals the connection between (irreducible) Markov chains and counting spanning trees/forests. One is a classical result, known as Markov Chain Tree Theorem(theorem 1), the other(theorem 2) is taken from [1], which shows Kemeny's constant can be expressed by the number of spanning forests. Spanning trees/forests are mainly combinatorial objects, while Markov chains are usually investigated in the context of probability theory, and thus such connections are often insightful and interesting.

Acknowledgement I would like to warmly thank Prof. Wu for his intriguing lecture on Graph and Networks during 2020 Fall. This note is also an extension of topics discussed during the lecture [1].

1 Introduction

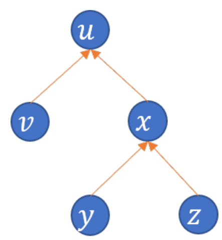
Let X be a irreducible Markov chain with finite state space $\mathcal{V}$ and transition matrix P. Let $\mathcal{A}_v$ be the set of *rooted spanning rooted trees* on $\mathcal{V}$ with root $v \in \mathcal{V}$. Write $\mathcal{A} := \bigcup_{v\in\mathcal{V}} \mathcal{A}_v$, i.e., $\mathcal{A}$ is the set of all rooted spanning trees on $\mathcal{V}$.

[1] A relevant material:
http://math.sjtu.edu.cn/faculty/ykwu/data/TeachingMaterial/MCT.pdf

We associate to each tree $T \in \mathcal{A}$ a weight via

$$w_P(T) := \prod_{(u,v)\in T} P(u,v),$$

where (u, v) runs over *directed* edges in T that goes towards the root. See fig. 1 for an example.



$T \in \mathcal{A}_u$

Figure 1: $w_P(T) = P(v, u)P(x, u)P(y, z)P(z, x)$. All arcs are oriented toward root.

Let

$$\Sigma_v := \sum_{T\in\mathcal{A}_v} w_P(T) \text{ and } \Sigma^{(1)} := \sum_{v\in\mathcal{V}} \Sigma_v.$$

In section 2, we shall prove

Theorem 1 (Markov chain tree theorem). *Let $\pi \in \mathbb{R}^{\mathcal{V}}$ with $\pi(v) := \Sigma_v/\Sigma^{(1)}$. Then π is the unique stationary distribution of X.*

We can generalize the weight function to a spanning forest of $\mathcal{V}$. Let $(T_1, T_2, \ldots, T_r)$ be a spanning forest of $\mathcal{V}$ with r rooted trees. Define

$$w_p(T_1, T_2, \ldots, T_r) := \prod_{i=1}^{r} w_p(T_i).$$

Analogously, write

$$\Sigma^{(r)} := \sum_{(T_1,\ldots,T_r)} w_p(T_1, T_2, \ldots, T_r),$$

where the sum is over all spanning forests consists of r rooted trees. In section 3, we shall prove that

Theorem 2. *Let κ be the Kemeny's constant of irreducible Markov chain X. Then*

$$\kappa = 1 + \frac{\Sigma^{(2)}}{\Sigma^{(1)}}.$$

2 Lifting the Markov chain to the spanning tree space

We shall present a natural way to 'lift' X to a Markov chains with state space $\mathcal{A}$, denoted by $\hat{X}$. Then we can prove theorem 1 by looking into the properties of $\hat{X}$. We first introduce projection of a Markov chain.

Projection of a Markov Chain. Let $X := (X_t)_{t\in\mathbb{N}}$ be a Markov Chain with state space $\mathcal{X}$ and transition matrix Q. Suppose that $f : \mathcal{X} \to \mathcal{Y}$ is a surjective map such that for all $y \in \mathcal{Y}$,

$$f(x_1) = f(x_2) \text{ implies } \sum_{x\in f^{-1}(y)} Q(x_1, x) = \sum_{x\in f^{-1}(y)} Q(x_2, x). \tag{1}$$

Define a matrix $Q' \in \mathbb{R}^{\mathcal{Y}\times\mathcal{Y}}$ by

$$Q'(y_1, y_2) := \sum_{x_2\in f^{-1}(y_2)} Q(x_1, x_2), \tag{2}$$

where x_1 is any preimage of y_1, i.e., $x_1 \in f^{-1}(y_1)$. Note that Q' is well-defined given that eq. (1) is satisfied. Let $Y_t := f(X_t)$. It is easy to check that $Y := (Y_t)_{t\in\mathbb{N}}$ is a Markov Chain with state space $\mathcal{Y}$ and transition matrix Q'. We say Y is a *projection* of X.

Intuitively, we expect that some properties of the original chain X can also be 'projected' to the projection chain Y. Indeed, stationary distribution is such an example:

Proposition 3. *Let Y be a projection of X. If $\pi \in \mathbb{R}^{\mathcal{X}}$ is a stationary distribution for X, then*

$$\mu(y) := \sum_{x\in f^{-1}(y)} \pi(x)$$

is a stationary distribution for Y.

Now we 'lift' X to a Markov Chain $\hat{X}$ with state space $\mathcal{A}$ and transition matrix $\hat{P}$, where $\hat{P}$ is defined as follows. Let $v \in V, T \in \mathcal{A}$, say, $T \in \mathcal{A}_u$. We obtain a new tree T', rooted at v, by adding the edge (u, v) to T and deleting the only out edge of v in T(see fig. 2). Then one simply set $\hat{P}(T, T') := P(u, v)$. If $T'' \in \mathcal{A}$ cannot be obtain from T in this way, set $\hat{P}(T, T'') = 0$.

Consider the mapping $\texttt{root} : \mathcal{A} \to \mathcal{V}$ that maps T to its root, that is, $\texttt{root}(T) := v$ for every $T \in \mathcal{A}_v$. We claim that X is a projection of $\hat{X}$, and

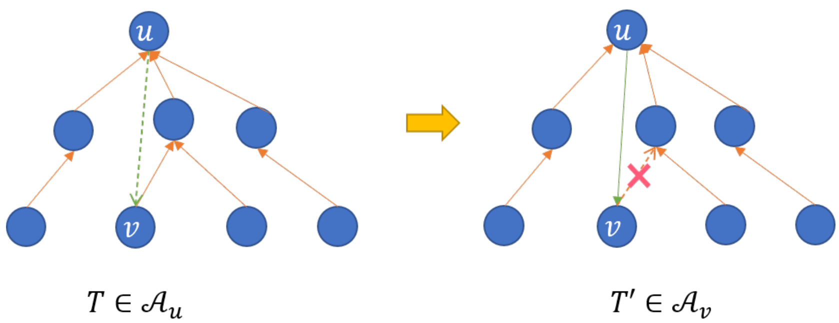


Figure 2: Transition of $\hat{X}$ with $\hat{P}(T, T') := P(u, v)$.

the projection mapping is `root`. Indeed, since the transition of $\hat{X}$ depends only on $\mathtt{root}(X_t)$, it is easy to verify eq. (1) and eq. (2) holds.

Now we turn to the proof of theorem 1.

Proof of theorem 1. $\hat{X}$ is also irreducible for we assume that X is irreducible. For $T \in \mathcal{A}$, let $\pi(T) := w_P(T)/\Sigma^{(1)}$. We claim that

Proposition 4. *π is the unique stationary distribution for $\hat{X}$.*

Proof. For any $T \in \mathcal{A}$, say $T \in \mathcal{A}_u$, we are to verify

$$\sum_{S \in \mathcal{A}} w_P(S)\hat{P}(S, T) = w_P(T).$$

Let $v_1, v_2, \ldots, v_\ell$ be the neighbors of u in T. For $i \in [\ell]$, Define

$$V_i := \{v \in \mathcal{V} : \text{there exists a path from } v \text{ to } v_i\}.$$

Note that if $\hat{P}(S, T) > 0$ and $S \neq T$, we have $S \in \mathcal{A}_{v_i}$ for some $i \in [\ell]$. Hence,

$$\sum_{S \in \mathcal{A}} w_P(S)\hat{P}(S, T) = w(T)\hat{P}(T, T) + \sum_{i=1}^{\ell} \sum_{S \in \mathcal{A}_{v_i}} w(S)\hat{P}(S, T) \tag{3}$$

For $S \in \mathcal{A}_{v_i}$, let $S = T - (v_i, u) + (u, x)$, where $x \in V_i$ is the predecessor of u in S, i.e., $(u, x) \in S$ (see fig. 3). Clearly, $w(S)\hat{P}(S, T) = w(S)P(v_i, u) = w(T)P(u, x)$ and $\hat{P}(T, T) = P(u, u)$. We point out that $S \mapsto x$ is a one-to-one correspondence. Therefore, together with eq. (3), we get

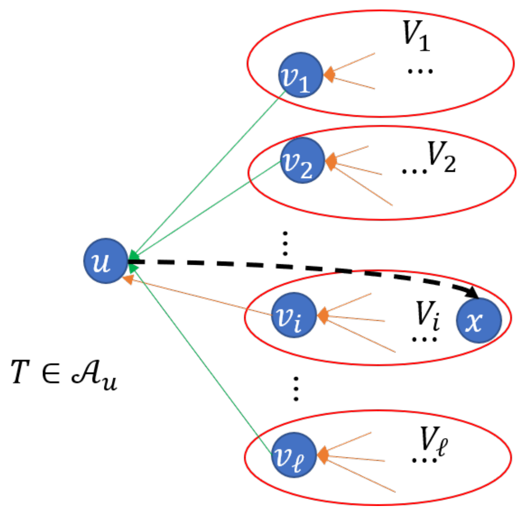


Figure 3: Tree S plus (v_i, u) equals tree T plus (u, x).

$$\sum_{S\in\mathcal{A}} w_P(S)\hat{P}(S,T) = w(T)P(u,u) + \sum_{i=1}^{\ell}\sum_{x\in V_i} w(T)P(u,x)$$

$$= w(T)\left[P(u,u) + \sum_{x\in\mathcal{V}\setminus\{u\}} P(u,x)\right] = w(T),$$

which is exactly what we set out to prove. □

We are happy to see that theorem 1 simply follows from proposition 4 and proposition 3. □

3 Kemeny's constant and counting spanning forests

As usual, we study an irreducible Markov Chain with state space $\mathcal{X}$ and stationary distribution π. For $x \in \mathcal{X}$, let τ_x be the *hitting time* of x.

3.1 Random walk with random Target

We are now interested in the quantity (for $a \in \mathcal{X}$):

$$\kappa(a) := \mathbf{E}_a[\tau_\pi] := \sum_{b\in\mathcal{X}} \mathbf{E}_a[\tau_b]\cdot\pi(b).$$

This quantity is the expected time of hitting a random target with distribution π.

Lemma 5 (Random Target Lemma). *The quantity $\kappa(a)$ does not depend on $a \in \mathcal{X}$.*

Proof. It suffices to show that κ is harmonic. Note that

$$(P\kappa)(a) = \mathbf{E}_a\left[\kappa(X_1)\right] = \sum_{x\in\mathcal{X}} \kappa(x)P(a,x) = \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{X}} \pi(b)\mathbf{E}_x\left[\tau_b\right]P(a,x). \quad (4)$$

If $b \neq a$,

$$\sum_{x\in\mathcal{X}} \mathbf{E}_x\left[\tau_b\right]P(a,x) = \mathbf{E}_a\left[\tau_b\right] - 1;$$

if $b = a$,

$$\sum_{x\in\mathcal{X}} \mathbf{E}_x\left[\tau_b\right]P(a,x) = \mathbf{E}_a\left[\tau_a^+\right] - 1 = \frac{1}{\pi(a)} - 1.$$

Hence, we rearrange eq. (4) as

$$\begin{aligned}
(P\kappa)(a) &= \sum_{b\in\mathcal{X}} \pi(b)\left[\sum_{x\in\mathcal{X}} \mathbf{E}_x\left[\tau_b\right]P(a,x)\right] \\
&= \sum_{b\in\mathcal{X}\setminus\{a\}} \pi(b)(\mathbf{E}_a\left[\tau_b\right] - 1) + \pi(a)(\frac{1}{\pi(a)} - 1) \\
&= \sum_{b\in\mathcal{X}} \pi(b)(\mathbf{E}_a\left[\tau_b\right] - 1) + 1 && \text{(since } \mathbf{E}_a\left[\tau_a\right] = 0\text{)} \\
&= \kappa(a).
\end{aligned}$$

This finishes the proof. □

According to the lemma above, the starting measure is of no significance for the quantity $\kappa(\cdot)$. Hence, we can define the *target time* of an irreducible chain by $\kappa := \kappa(a)$, where $a \in \mathcal{X}$ is arbitrary. Or equivalently,

$$\kappa := \mathbf{E}_\pi\left[\tau_\pi\right] := \sum_{a,b\in\mathcal{X}} \mathbf{E}_a\left[\tau_b\right]\pi(a)\pi(b).$$

Intuitively, κ is the expected time of going to a random target from a random starting location. Thus, we deem that κ measure the connectivity of the network to some degree. κ is also known as **Kemeny's constant**.

3.2 Relating Kemeny's constant and spanning forests

For simplicity, write $m(u, v) := \mathbf{E}_u[\tau_v]$ for every $(u, v) \in \mathcal{V} \times \mathcal{V}$. For $T \in \mathcal{A}_v$, let $\texttt{last}(T, u)$ be the last vertex before v in the path from u to v in T, as is shown in fig. 4.

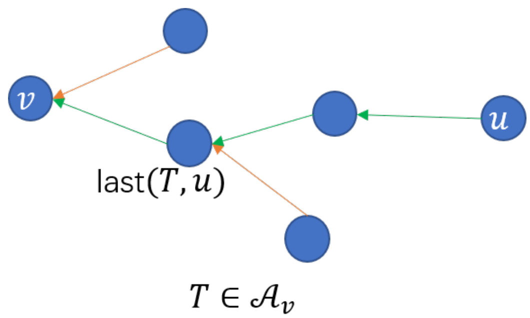


Figure 4: Definition of $\texttt{last}(T, u)$.

In order to prove theorem 2, we draw on the following theorem without presenting a proof.

Theorem 6 (Markov chain tree formula for mean hitting times [1]). *Let P be a transition matrix for an irreducible chain. For each $u \neq v$,*

$$m(u, v) = \frac{\Sigma_{uv}}{\Sigma_v},$$

where

$$\Sigma_{uv} := \sum_{T \in \mathcal{A}_v} \frac{w_P(T)}{P(\texttt{last}(T, u), v)}.$$

Now we are ready to prove theorem 2.

Proof of theorem 2. By omitting the arc $(\texttt{last}(T, u), v)$ in T, we get a spanning forrest (T_1, T_2) from T, where T_1 is the subtree with root $\texttt{last}(T, u)$, $T_2 = T \setminus T_1$. The map $T \mapsto (T_1, T_2)$ is a bijection from $\mathcal{A}_v$ to the spanning forests

$$\mathcal{F}_{uv} := \{(T_1, T_2) : u \in T_1 \text{ and } \texttt{root}(T_2) = v\}.$$

Note that $w_p(T)/P(\texttt{last}(T, u), v) = w_P(T_1, T_2)$, and thus

$$\Sigma_{uv} = \sum_{(T_1, T_2) \in \mathcal{F}_{uv}} w_P(T_1, T_2).$$

Observe that for fixed u, $\{\mathcal{F}_{uv} : v \in \mathcal{X} \setminus \{u\}\}$ form a partition of all 2-component spanning forests. Hence,

$$\Sigma^{(2)} = \sum_{v \in \mathcal{X} \setminus \{u\}} \Sigma_{uv}. \tag{5}$$

Since $m(u,u) = \frac{1}{\pi(u)}$, by theorem 6 we have

$$\kappa = 1 + \sum_{v \in \mathcal{X} \setminus \{u\}} \pi(b) m(u,v) = 1 + \sum_{v \in \mathcal{X} \setminus \{u\}} \pi(v) \frac{\Sigma_{uv}}{\Sigma_v}. \tag{6}$$

According to theorem 1, $\pi(v) = \Sigma_v / \Sigma^{(1)}$. Plugging this into eq. (6) yields

$$\kappa = 1 + \sum_{v \in \mathcal{X} \setminus \{u\}} \frac{\Sigma_{uv}}{\Sigma^{(1)}} = 1 + \frac{\Sigma^{(2)}}{\Sigma^{(1)}},$$

where the last step follows from eq. (5). □

References

[1] Jim Pitman and Wenpin Tang. Tree formulas, mean first passage times and kemenys constant of a markov chain. *Bernoulli*, 24(3):19421972, Aug 2018. 1, 7